A Novel Optimization based Algorithm to Hide Sensitive Item-sets through Sanitization Approach

T.Satyanarayana Murthy, N.P.Gopalan
National Institute of Technology /CA, Tiruchirapalli, 620015, INDIA.
Email: murthyteki@gmail.com,npgopalan@nitt.edu

Sasidhar Gunturu
Vignan University/CSE, Guntur, INDIA
Email: g.sasidhar1997@gmail.com

Abstract—Association rule hiding an important issue in recent years due to the development of privacy preserving data mining techniques for hiding the association rules. One of the mostly used techniques to hide association rules is the sanitization of the database. In this paper, a novel algorithm MPSO2DT has been proposed based on the Particle Swarm Optimization (PSO) in order to reduce the side effects. The aim is to reduce the side effects such as Sensitive item-set hiding failure, Non-sensitive misses, extra item-set generations and Database dissimilarities along with the reduction of running time and complexities through transaction deletion.

Index Terms—Association Rule Hiding, Particle Swarm Opti-mization, Hiding Failure, Non-Sensitive Misses. Database Dissimi-larity

I. INTRODUCTION

A database may contain astronomically immense amount of data. Discovering useful information and associations from data known as Knowledge Discovery in Database (KDD) has become facile with magnification of information technology. Sharing of such data may lead to leakage of some sensitive in-formation and may become a threat to privacy. The KDD techniques are classified into Association rules and it was proposed by Agrawal and Srikant in 1993 & 1994 [1,2] Chen et al [3] in 1996. An incipient field in data mining named as Privacy Preserving Data Mining (PPDM) was introduced by Agrawal et al.,2006 [4] while endeavoring to obtain patterns that are useful without compromising the sensitive data privacy. The competitors can utilize mining results in areas of business and other fields to discover relationship between transactions and become a security threat to one's business, thus association rule hiding/data sanitization in order to allow analyzing of data at the same time hiding sensitive rules was introduced. There are various side effects of sanitizing data. Some sensitive data may be still visible in the sanitized database causing hiding failure, at the same time some non-sensitive data that was supposed to be visible for analysis in the sanitized database may be hidden. Also, some extra rules may be generated in the modified database. The ordinary algorithms used for sanitization are not effective in minimizing the side effects of sanitization which is linked with the selection of correct transaction for modification. The traditional PPDM algorithms Identify applicable funding agency here. If none, delete this. are not efficient in searching the near optimal solutions in short time. The evolutionary algorithms are better while considering the NP-hard problems. Genetic algorithm [5,6] is one of the widely used and useful evolutionary algorithm. This population-based algorithm reach a near optimal solution by performing crossover, mutation and a selection operation. This can be applied to various discrete or continuous problems. Particle swarm optimization (PSO) is a population-based algorithm invented by Kennedy and Albert (1995) [7,8] that follows a path towards an optimal solution similar to a flock of birds finding better food resources. In PSO the possible solutions are considered as different particles which are evaluated using a predefined fitness function. The particle speed determines the location and the speeds are up-dated in each iteration. This updation was performed by using two functions global best (gbest) and personal best(pbest). Cuckoo optimization (Afshari, M. H. Dehkordi, M. N. Akbari, M. (2016) was another approach for hiding data using distortion technique. This method is based on the survival technique of bird cuckoo. This bird never builds its own nest and puts its eggs in other bird's nest and still and makes them participate in its survival. This algorithm introduces three fitness functions to reduce side effects and a better immigration function in order to escape from local optimum. Ant colony (Doringo and Gambardella 2016) an advanced optimization technique used for association rule hiding, where it uses various pheromone updating rules and state transaction rules as an extension to the ordinary AS (Colorni et al. And Doringo et al.) so

that efficiency is increased. In this method the concept used is the deposition of pheromones by the ants in the path leading to help them find shortest path between their location and nest. The pheromone density is updated at each iteration and a designed heuristic function are considered for selecting route. This paper contains 10 sections. Section 2 contains the literature survey about the various evolutionary algorithms. Section 3 specifies the preliminaries and basics of association rule hiding. Section 4 describe the flow of the proposed algorithm, Section 5 describes the performance parameters of association rule hiding. Section 6 and Section 7 states the problem and the proposed algorithm. Section 8 contains experimental results and the last section gives the conclusion and future work.

II. LITERATURE SURVEY

Atallah et al.,[9] proposed a Disclosure limitation of sensitive rules uses cyclic based approach for handled the sensitive rules.Three algorithms 1.a, 1.b, 2.a were presented by Dasseni et al. [10] in which 1.a, 1.b was based on the concept of increasing support of the antecedent of the rule and thus reducing the confidence of the rule. The algorithm 2.a was based on the concept of decreasing the support of the itemset of rule generated. Some algorithms similar to previous three Confidence reduction (CR), CR2, and Generating Itemset Hiding (GIH) were proposed by Saygin et al. (2001). The work of Dasseni et al. (2001) was extended by Verykios et al.(2004b) [11]and introduced algorithm 2.b in order to hide generating itemset of sensitive rules.Wang et al., 2007[12] with the aim of hiding predictive association rules, proposed an algorithm Decrease Support and Confidence (DSC) algorithm. The constraint Satisfaction problem proposed by Menon et al. (2005) was extended by Menon and Sarkar (2008) to reduce Not-To-Hide rules and number of sanitized transactions. B Keshava Murty proposed a Privacy preserving association rule mining over distributed databases using genetic algorithm for hiding the sensitive rules. 2014: Lin et al. (2014a, 2014b) [13,14] in order to select transactions for hiding itemset used Genetic algorithms for the first time in which Lin et al (2014a) contain Compact Pre-large GA-based algorithm to Delete Transactions(cpGA2DT) that deletes the transactions specified and Lin et al., (2014b) contains an algorithm proposed to produce and insert new transactions. On the basis of PISA platform (Bleuler et al., 2003) an Evolution based Multi-objective Optimization base Rule Hiding (EMO-RH) algorithm was proposed by Cheng et al.. Lin et al., (2015) [15,16,17] proposed algorithms Simple Genetic Algorithm to Delete Transactions (sGA2DT) and Pre-large Genetic Algorithm to Delete transactions. Particle swarm optimization based algorithm to Delete Transaction (PSO2DT) was proposed by Lin et al. [18,19] (2016) with less parameters and also to find the number of transactions to be deleted to minimize side effects which was earlier done manually. Cuckoo Optimization Algorithm for Association Rule Hiding (COA4ARH) by Afsari et al. (2016) [20] was

proposed to hide sensitive rules using Cuckoo Optimization Algorithm (Yang Deb,2009). The algorithm varies in the concept of performing a pre-processing operation that contain 2 phases in order to reduce number of loops and time to get an optimal solution. 2017: Telikani Shahbahrami (2017) [21] improved Max Min solution (Moustakides Verykios, 2006, 2008) with 2 heuristics to hide association rules. This algorithm is named DCR (Decrease the confidence of Rule). J. M.-T. Wu et al.: Ant Colony System Sanitization approach for hiding sensitive item sets., (2017). JIMMY MING-TAI WU [22] proposed an Ant Colony Optimization based algorithm for association rule hiding. T. Satyanarayana Murthy et al., [23,24,25,26,27] proposed meta heuristic based algorithms for better way of association rule hiding.

In this paper we extend PSO2DT algorithm Lin et al. (2016). The proposed Modified PSO2DT checks for the possibility to reduce the number of transactions to be deleted and thus reduce dissimilarity in the sanitized and the original database.

III. BASIC PRELIMINARIES FOR ASSOCIATION RULE HIDING

A. Definition 1

Let A be an itemset then, the number of appearances of A in the database Db is its support count. The support of an itemset is the total number of transactions containing that itemset. Let the number of transactions in database be N and the Minimum support threshold .Let the database contain 10 transactions with Minimum support count= $(10 \times 0.4) = 4$.

B. Definition 2

If support count of any item is more than or equal to MSC, then it is a frequent item. The set of itemset that is sensitive is denoted by SIs and it is a subset of frequent set FIs. The sensitive itemset is decided by the user or expert. Here if MSC=4, all the item sets in the original database with item count more than or equal to 4 are frequent item sets.

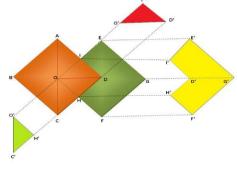### C. Side Effects in Association Rule Hiding Process

Fig.1. ARH Process

Fig.1 shows the relationship between the side effects of Association rule hiding process. ABCD represents original dataset and EOHD is the sanitized dataset i.e. the

output.

The various side effects are described thereafter. In ABCD, ABC is the non-frequent itemset (~FIs) i.e. Support less than Min_ Support_ threshold.

ACD represent set of item sets that are frequent (FIs) in the database i.e., Support M in Support threshold. AOD is the set of sensitive data items (SIs)specified by user or expert; OCD is the non-sensitive item sets (~SIs).

D. Definition 3:

OCH is the data items which were supposed to be present in sanitized database but a miss occurs and it remains hidden. This is called Non-sensitive Miss (NSM) and represented as β, where $\beta = |\sim SIs-FIs*|$.

E. Definition 4:

In EOFG (Sanitized dataset), IOD is the sensitive dataset of original dataset that was supposed to be hidden in sanitized dataset but the sanitization algorithm could not hide these items. Thus, a Sensitive Hidden Failure (SHF) occurs. This is represented as α where $\alpha = |SIs \cap FIs|$.

F. Definition 5:

EIDHFG is another side effect where the extra or ghost rules are generated due to sanitization process. This is called ERG represented as γ EIDHFG is the extra or ghost rules that are generated as a side effect of sanitization process. This is called ERG represented as γ ,where $\gamma = |FIs* - FIs|$.

G. Definition 6:

Let the original database be Db and the sanitized database be Db* and then the similarity in Db* and Db must be maintained high, with value not less than 90, which is the currently achievable value in most of the algorithms. Database similarity = $|Db*|/ |Db|$. A high value of database similarity indicates that the sanitized database still reflects the original database structure and the deletion and insertions did not affect the overall database to larger extend.

IV. MODIFIED PSO2DT FLOWCHART

The Fig.2 shows the flow chart indicating the main steps involved in the proposed MPSO2DT algorithm. The sensitive data item sets and original database are given as input to the algorithm. A projected database is generated from the original database. The transactions from the projected database are selected randomly to assign it to each particle. Now for each particle fitness values are calculated and then the personal best (pbest) and the global best(gbest) of all the particles are updated in each iteration. In the process of finding gbest a decision is to be made. If the pbest values are same for more than one particle and it is the lowest pbest value then the particle with lowest number of transactions to be deleted is taken as gbest. After the completion of all iterations, the last gbest value is selected and the transactions present in that particle are deleted.

V. PERFORMANCE PARAMETERS OF ASSOCIATION RULE HIDING

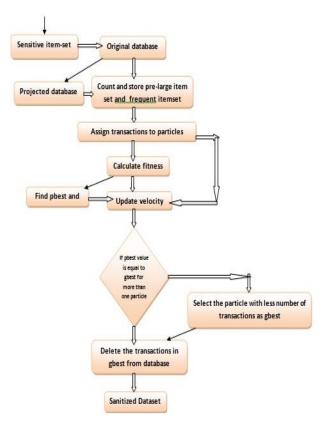Some of the most predominant performance parameters of association rule hiding are

Fig.2. MPSO2DT Flow Chart

a) Hiding Failure: Some sensitive rules fail to be hidden in the sanitized database.

b) Lost rules: Some of the frequent item sets that are non-sensitive remain hidden in sanitized database. This leads to unnecessary hiding of data.

c) Ghost rules: Ghost rules are the additional rules which are generated because of sanitization process.

d) Database dissimilarity: The ratio between the sanitized database Db* and the original database Db must be kept minimum i.e. the deletion or change in the transactions of the original database must be minimum. Database similarity ratio must be more than 90

e) Computational complexity: The computational time complexity must be minimum i.e. the sanitization algorithm must be able to produce the resultant database with lesser time than other algorithms. The lesser the complexity, more efficient is the algorithm.

f) Accuracy: The degradation of data accuracy leads to resulting of the knowledge extracted from the sanitized database to be useless. The accuracy is inversely proportional to the dissimilarity in the databases.

VI. Problem Definition

Consider the database Db associates with a set of tuple T (T1, T2....Tn). Each tuple associates with a set of attributes like A (A1, A2, A3...An). The hiding of association rule by sanitization of the original database is followed by the side effects Sensitive data hiding failure, Non-sensitive Misses and the Extra item sets being generated. Thus, to reduce the side effect PSO based algorithm PSO2DT deletes the transactions to obtain sanitized database. At the selection of gbest values in this algorithm whenever there are 2 particles with same lowest values that can be considered for gbest election, the existing algorithm selects the last particle that contain the lowest value. This is not the optimal solution in all situations, as the transactions to be deleted in each particle may vary. To address this problem, a modified PSO2DT algorithm is proposed.

VII. Proposed Alogorithm

Input : A projected database Db', Frequent item sets FIs , Set of sensitive dataset Sis ,Minimum support threshold δ ,Number of particle N

Output: Sanitized database Db*

1. Calculate particle size
$$m = \left\lceil \frac{Max\_Support(SIs) - \delta \times |Db|}{1 - \delta} \right\rceil$$
2. For i = 1 to N
3.   For j=1 to m do
4. $particle_i(x) \leftarrow particle_j(x) \cup (rand(Db'), Null)$
5. While end condition not satisfied
6. For i = 1 to N
7. $fitness(particle_i(x)) = w_1\alpha + w_2\beta + w_3\gamma$
8.   if $fitness(particle_i(x)) \leq pbest_i(x)$ then
9. $pbest_i(x) \leftarrow particle_i(x)$
10.   If $pbest_i(x) \leq gbest$ then
11. $gbest \leftarrow pbest_i(x)$
12. Count=0
13. For i=1 to N
14.   If $fitness(particle_i(x)) == gbest$
15.   Count++;
16. If count>1
17.   For i=1 to N
18.   For j=1 to m
19.   Count=Number of not null transactions in $particle_i(x)$
20.   if min>count
21.   Min=count
22. $gbest \leftarrow fitness(particle_i(x))$
23. For i=1 to N
24. Update particles
25. Set $x \leftarrow x + 1$
26. Delete transactions present in $gbest$
27. Return Db^*

This algorithm performs the following steps Let maximum number of transactions to be deleted be m then m = CEIL ((Max Support (SIs) – δ |Db|)/ (1 − δ)) this is the first step in the algorithm where SIs is the set of sensitive itemset. The number of transactions to be deleted is less than or equal to m. Thus, the particle size is taken as m with initialization with transactions present in the projected database and also null values represented as 0, both chosen randomly. Finding pre-large itemset the concept of pre-large enables the lesser access of database since once the pre-large itemset are identified only these items need to checked for ERG. Next the fitness value for each of the particles using the existing formula fitness(particle$_i$(x)) = w$_1$+w$_2$+w$_3$ Where w1, w2, w3 are the coefficients of the Sensitive Hidden failure, Non-sensitive misses and Extra Rules generated respectively. The values of these are selected as per the users requirement and importance of each side effect. If the w1 value is set high it indicates that the expert or user wish to hide maximum sensitive data. If w2 is set high then it indicates the preserving of maximum non-sensitive data for deriving decisions is more important. Also, if w3 is high, it indicated the reduction of extra rules is the major requirement. Let us consider w1=0.8, w2=0.1 and w3=0.1.

Next step is finding the pbest value. The pbest value is the solution of a particle which is its personal or individual best value (according to the fitness function) so far. Next step is finding the value of gbest The gbest or the global best calculation involves 2 step Select the smallest pbest value as gbest. If more than one particle have same pbest value, considering for gbest then instead of selecting the last particle as gbest, the proposed algorithm selects the particle with less number of transactions as the gbest. This finds the possibility of reducing the no. of transactions to be deleted to get the sanitized database. Deleting transactions from the gbest is the final step to get sanitized database Db* by reducing side effects.

VIII. Experimental Results

The comparison between the existing PSO2DT algorithm and the modified algorithm was done by conducting various experiments. The algorithms in the experiments were implemented in Python language on a computer of features windows 10 * 64 equipped with processor Intel 1.8GHz corei3 and RAM 4GB. Three datasets chess[1], mushroom[2] which are dense and food-mart[3] , a sparse dataset was taken was taken for this experiment. In this section the various factors such as Sensitive hiding failure, Non sensitive misses, Extra rules generated and database dissimilarities of the GA based algorithm cpGA2DT, existing PSO2DT algorithm and the modified algorithm MPSO2DT are compared. Each of the performance parameter values are noted by varying the MST and the result is represented in the form of bar graph.

A. Side effects evaluation

The comparison between the original and the sanitized database can be done by comparing the side effects. PSO2DT and MPSO2DT both showed a higher ability to hide sensitive dataset. At the same time these algorithms generated many N-S-M and E-R-G in some datasets. In this evaluation the existing and the modified algorithm gives the similar result as there are no variation in the steps involved till the side effects generation.

B. NSM

The NSM is to find how many Non-sensitive frequent item sets got hidden during the sanitization process which can be formulated as: NSM = |FIs−SIs−FIs∗|/|FIs−SIs|.
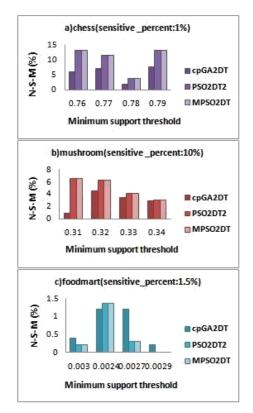

Fig.3. N-S-M Results

The denominator in the above formula |FIs − SIs| is the items in original database that are non-sensitive and the numerator |FIs−SIs−FIs∗| indicate the hidden non-sensitive item sets. In Fig.3 it can be observed that the MPSO2DT algorithm generates the same amount of NSM as that of existing algorithm. The N-S-M is less for the modified algorithm when compared to algorithm based on GA in most of the cases of sparse dataset such as food-mart as shown in fig 3.c along with a nearly zero S-H-F on an average as shown in fig 4.c. But it can be observed from fig 3.a and fig 3.b that the N-S-M in a dense dataset such as chess and mushroom is more for the modified algorithm as it is directly connected with the no. of S-H-F .When the maximum Hiding failure is reduced, there are more chances that N-S-M also increase. Thus, in a dense database when almost the S-H-F is zero, it produce N-S-M.

C. S-H-F

The S-H-F is an evaluation metric to find the extend of the Sensitive data item sets that could not be hidden in Db*. It is given by S −H −F = |SIs∗|/ |SIs| Where SIs* is the sensitive data item-sets that are still present in sanitized database which was supposed to be hidden. And SIs* is a superset of SIs* which is the sensitive dataset of original database. Fig.4 a,b,c shows the results obtained when the various MSC effected the no. of S-H-F. In Fig.4.a for dense dataset chess the modified algorithm produce no S-H-F side effects whereas the GA based algorithm cpGA2DT generates above 40% S-H-F for most of the minimum support threshold values. Also, for support values of 0.33 and 0.44 the mushroom dataset has no S-H-F and has a very low percentage of S-HF as compared to GA based algorithms for minimum support threshold 0.31 and 0.32 shown in fig 4.b. At the same time for dataset food-mart fig 4.c, the modified algorithm performance is better than the referenced GA based algorithm and is almost zero for most of the values of minimum threshold.
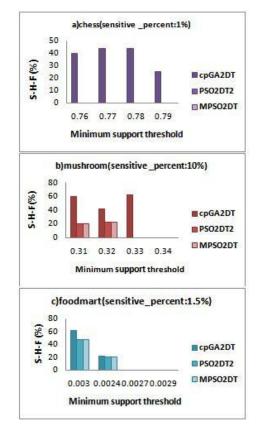

Fig.4. S-H-F Results

D. Fitness

The fitness value of each particle is calculated as the sum of products of the number of each side effect with their corresponding weights assigned according to the users preference. The weight of side effect which has to be solved mainly is kept high i.e. if sensitive item sets hiding is our foremost goal then the value of w1 must be

higher than the other two. Here for dense datasets chess and mushroom, the weights w1, w2, w3 are taken as 0.98,0.01,0.01 respectively and for sparse dataset food mart the values are 0.8,0.1,0.1. The fitness values obtained for each dataset with different minimum support thresholds are given in Fig.5 a, b, c. from Fig.5 indicate that the modified algorithm has better fitness values than the GA based algorithm. The modified algorithm is designed to hide all the sensitive itemset, thus it has low fitness values compared to GA based algorithm. In the fig 5.a and 5.b fitness values are due to the N-S-M in the dense datasets. Meanwhile in a sparse dataset fig 5.c there is no much difference between the fitness values from other algorithms. Thus, it is observed that the modified algorithm shows a performance similar to the existing algorithm and there is no degradation in efficiency.
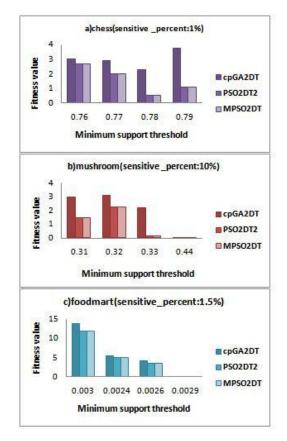

Fig.5.Fitness Details

E. *Database similarity*

The database similarity is the key concept of this paper. The minimum difference between a sanitized database and an original database indicates that the transaction selection for the generation of sanitized database was the best. The Database similarity is given as Database similarity = |Db*|=|Db| Where —Db* — is the no. of transactions in the sanitized database and —Db— is the total no. of transactions in the original database. The main aim of the modified algorithm is the reduction of no. of transactions to be deleted in the sanitization process and thus avoid deleting unnecessary trans-actions. From the

fig 6 it can be observed that all the compared algorithms have data similarity above 90%. The MPSO2DT algorithm perform better than the existing algorithm as it always searches for the condition where the transaction selected for the deletion be kept minimum. In a dense dataset the difference is more than the sparse dataset as shown in Fig.6 a, b, c. Thus the modified algorithm reduces the transactions to be deleted, to find a better sanitized database with less no. of side effects. Though the side effects and fitness values are not affected by the modified algorithm, it definitely improves the performance of the algorithm by selecting more appropriate transactions without increasing the execution time.
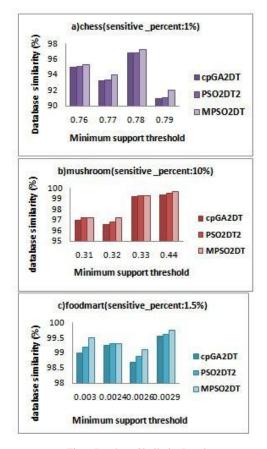

Fig.6. Database Similarity Results

IX. CONCLUSION

Many algorithms exist in order to generate a sanitized database by the method of transaction deletion. A particle swarm optimization PSO2DT algorithm is one of the effective algorithms which perform better than other Genetic algorithm based techniques. In this paper, a modified PSO2DT algorithm MPSO2DT is proposed which perform better than the existing algorithm. In MPSO2DT, the transactions to be deleted is reduced and thus reducing the data dissimilarity without affecting the other parameters.

REFERENCES

[1] Rakesh Agrawal, Tomasz Imielinski, Arun Swami," Mining association rules between sets of items in large databases". *ACM SIGMOD international conference on Management of data SIGMOD*, pp. 207,1993.

[2] Agrawal,R.,Srikant,R., a.QuestSyntheticDataGenerator . IBM AlmadenRe- searchCenter ⟨http://www.Almaden.ibm.com/cs/quest/syndata.html⟩,1994.

[3] Chen, M.S.,Han,J.,Yu,P.S.,"Datamining:An overview from a database per spective". *IEEE Trans. Knowl. DataEng*.8(6),866–883,1996.

[4] Aggarwal,C.C.,Pei,J.,Zhang,B.,."On privacy preservation against adversarial data mining",*ACM SIGKDD International Conference on Knowledge Dis- covery and Data Mining*,pp.510–516, 2006.

[5] Goldberg, D.E., "Genetic Algorithms in Search,Optimization and Machine Learning". Addison-Wesley Longman PublishingCo.,Inc,Boston,MA,USA,2002.

[6] Goldberg, David ," The Design of Innovation: Lessons from and for Competent Genetic Algorithms", *Norwell, MA: Kluwer Academic Publishers*,2002.

[7] Kennedy,J.,Eberhart,R," Particle swarm optimization ",*IEEE International Conference on Neural Networks*,pp.19421948.

[8] Kennedy,J.,Eberhart,R.,."A discrete binary version of particle swarm algorithm". *In IEEE International Conference on Systems,Man,and Cybernetics*, pp. 41044108997.

[9] Atallah, M. , Bertino, E. , Elmagarmid, A. , Ibrahim, M. , Verykios, V., "Disclosure limitation of sensitive rules", *In IEEE knowledge and data engineering exchange*, pp. 45–52,1999.

[10] Dasseni,E.,Verykios,V.S.,Elmagarmid,A.K.,Bertino,E.,"H iding association rules by using confidence and support", *International Workshop on Information Hiding,*pp.369–383,2001.

[11] Verykios, V. S. , Elmagarmid, A. K. , Bertino, E. , Saygin, Y. , Dasseni, E.," Association rule hiding", *IEEE Transactions on Knowledge and Data Engineering*, pp. 434–447,2004 .

[12] Wang, S.-L. , Parikh, B. , & Jafari, A. "Hiding informative association rule sets". *Expert Systems with Applications*, 33 , pp-316–323,2007.

[13] B.Kesava Murthy ,Asad M.Khan," Privacy preserving association rule mining over distributed databases using genetic algorithm", *Neural Computing and Application* ,2013.

[14] C.-W. Lin, T.-P. Hong, C.-C. Chang, and S.-L. Wang, "A greedy-based approach for hiding sensitive itemsets by transaction insertion," *J. Inf. Hiding Multimedia Signal Process*., vol. 4, no. 4, pp. 201–227, 2013.

[15] C.-W. Lin, B. Zhang, K.-T. Yang, and T.-P. Hong, "Efficiently hiding sensitive itemsets with transaction deletion based on genetic algorithms, "Sci. World J., vol. 2014, pp. 1–13, Sep. 2014.

[16] C.-W. Lin, T.-P. Hong, K.-T. Yang, and S.-L. Wang, "The GA-based algorithms for optimizing hiding sensitive itemsets through transaction deletion," Appl. Intell., vol. 42, no. 2, pp. 210–230, 2015.

[17] J. C.-W. Lin, L. Yang, P. Fournier–Viger, M.-T. Wu, T.-P. Hong, and L. S.-L. Wang, "A swarm-based approach to mine high-utility itemsets," i*n Proc. Multidisciplinary Social Netw. Res.*, pp. 572–581,2015.

[18] J. C.-W. Lin, Q. Liu, P. Fournier-Viger, T.-P. Hong, M. Voznak, and J. Zhan,"*A sanitization approach for hiding sensitive itemsets based on particle swarm optimization," Eng. Appl. Artif. Intell.*, vol. 53, pp. 1–18, Apr. 2016.

[19] J.C.W. Lin, Q. Liu, P. Fournier-Viger, T.P. Hong, M. Voznak, J. Zhan,"A sanitization approach for hiding sensitive itemsets based on particle swarm optimization", *Engineering Applications of Artificial Intelligence*,pp 1-18, 2016.

[20] Mahtab Hossein Afshari ,M.N.Dehkordi,M.Akbari" Association rule hiding using cuckoo optimization algorithm" *Expert Systems With Applications,* vol 64, pp 340–351,2016.

[21] Telikani, A., Shahbahrami, A., Optimizing association rule hiding using combination of border and heuristic approaches. Applied Intelligence 47, 544–557,2017.

[22] Jimmy Ming-Tai Wu,Justin Zhan and Jerry Chui Wei Lin "Ant Colony System Sanitization Approach to Hiding Sensitive Itemsets" *in IEEEAccess*,2017.

[23] N.P.Gopalan, T.Satyanarayana Murthy, Yalla Venkateswarlu, "Hiding Critical Transactions using Un-realization Approach", *IJPAM*, Vol 118,No.7 ,629-633,2018.

[24] T.Satyanarayana Murthy, N.P.Gopalan, "A Novel Algorithm for Association Rule Hiding", *International Journal of Information Engineering and Electronic Business (IJIEEB)*, Vol.10, No.3, pp. 45-50, 2018. DOI: 10.5815/ijieeb.2018.03.06

[25] T.Satyanarayana Murthy, N.P.Gopalan, Yalla Venkateswarlu," An efficient method for hiding association rules with additional parameter metrics ",*IJPAM* ,Vol 118,No.7,285-290,2018.

[26] T.Satyanarayana Murthy, N.P.Gopalan, , "Association rule hiding using chemical reaction optimization",*SCOPRO 2017 Conference,IIT Bhubaneswar*,2017,(Accepted).

[27] T.Satyanarayana Murthy.,"Privacy Preserving for expertise data using K-anonymity technique to advise the farmers", *International Journal of Electrical, Electronics and Data Communication*, Volume-1, Issue-10, 2013.

Authors' Profiles

N.P.Gopalan Professor at Department of Computer Applications, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India. He obtained his Ph.D. from Indian Institute of Science, Bangalore, India. His research interests are in Data Mining, Distributed Computing, Cellular Automata, Theoretical Computer Science, Image Processing and Machine Intelligence.

T.Satyanarayana Murthy Department of Computer Applications, National Institute of Technology, Tiruchirappalli. He obtained his B. Tech. in Information Technology from Jawaharlal Nehru Technological University, Hyderabad, Andhra Pradesh, India in 2006 and M. Tech. in Computer Science Engineering from Andhra University, Visakhapatnam, Andhra Pradesh, India in 2010. His current research interests include Data Mining, Privacy Preserving, CBIR, Big Data, Soft Computing.

Sasidhar Gunturu Pursuing B. Tech in the Stream of Computer Science and Engineering, Vignan's University, Vadlamudi, Guntur. His areas of interest are Data Analytics and Privacy Preserving Data Mining Issues.